

Un exemple de modélisation mathématique pour la fiabilité : le brûlage

Philippe Soulier^(*)

Résumé. Le brûlage est une technique industrielle de test ayant pour but d'éliminer les éléments faibles de la production, voire de renforcer la qualité de la production en diminuant le taux de défaillance après mise en service. Cet article présente des outils et des résultats mathématiques applicables à cette pratique. Ces outils proviennent essentiellement de la théorie des probabilités.

1. Introduction

Le brûlage est une méthode de fiabilité destinée à éliminer des éléments faibles ou défectueux d'une population de composants ou de systèmes industriels. Chaque composant individuel ou chaque système assemblé est mis en service dans des conditions normales ou extrêmes d'utilisation pendant une certaine durée. Les composants ayant survécu au test sont mis en service et les composants n'ayant pas survécu sont éliminés ou réparés et soumis à nouveau au test. Après assemblage, on peut procéder à un nouveau test. Cette procédure peut être, par exemple, appliquée à des ampoules. Dans le cas de systèmes complexes, le brûlage est assimilable à une forme de rodage. Le manuel d'installation des centraux téléphoniques d'AT&T décrit la procédure de rodage de ses centraux. La problématique du brûlage est donc au premier abord de nature physique. La nature physique ou chimique des composants ou systèmes testés déterminera l'intérêt du brûlage et, le cas échéant, ses modalités.

Que peuvent donc apporter les mathématiques au traitement d'un tel problème, mis à part les outils qu'utilise la physique ? Rien ! En fait, pour aborder ce problème mathématiquement, il faut le reposer d'un point de vue différent. La description initiale du problème est idéale. En réalité, le brûlage n'a pas de justification physique rigoureuse, et sa pratique relève de considérations souvent psychologiques. En premier lieu, elle traduit l'incertitude du producteur quant à la qualité, à l'homogénéité et à la durée de vie théorique de sa production. Cette incertitude rend naturelle la modélisation mathématique du problème dans les termes de la théorie des probabilités. La durée de survie d'un composant est une quantité aléatoire ayant une certaine loi de probabilité ou distribution.

Les deux problèmes mathématiques que pose le brûlage sont de déterminer pour quel type de distribution de survie le brûlage est une procédure pertinente et, pour une distribution donnée, de déterminer une durée optimale de brûlage en fonction de certains critères prédéfinis.

(*) Département de Mathématiques, Université d'Évry Val d'Essonne, 91025 Évry cedex, France.

2. Quelques rappels de théorie des probabilités

Dans la terminologie mathématique, le résultat d'un phénomène aléatoire est appelé variable aléatoire. Formellement, c'est une application d'un espace Ω dans l'espace des résultats possibles. Pour l'étude des durées de survie, on considère des variables aléatoires, notées T , à valeurs dans \mathbf{R}_+ . L'espace Ω reste indéterminé, et l'application T ne peut être exprimée explicitement : on ne peut prévoir le résultat de l'expérience. Ce que l'on connaît avant expérience, c'est la probabilité de la réalisation de la variable dans telle ou telle modalité. L'exemple élémentaire est celui d'un lancer de dé. On ne peut prévoir mécaniquement le résultat du lancer de dé, mais on sait *a priori* que chaque face du dé est équiprobable. Dans le contexte présent, on ne sait pas à l'avance quelle sera la durée de vie d'un composant individuel, mais l'on connaît théoriquement la probabilité que sa durée de vie dépasse une certaine valeur.

Rappelons maintenant le sens mathématique de probabilité. Une probabilité est une application définie sur un ensemble⁽¹⁾ de parties de Ω , appelées les événements, ayant les propriétés suivantes :

- $P(\Omega) = 1$; Ω est appelé l'événement certain.
- Si A et B sont des événements disjoints, $P(A \cup B) = P(A) + P(B)$. Cette propriété doit être aussi valide pour des collections dénombrables d'événements.

Dans l'étude probabiliste des durées de vie, les objets mathématiques d'intérêt sont les suivants :

- La fonction de répartition F de la variable T , définie sur \mathbf{R}_+ par :

$$F(t) = P(T \leq t).$$

$F(t)$ est la probabilité que le composant tombe en panne avant t . F est une fonction croissante, égale à 0 en zero et 1 en l'infini. La fonction de répartition caractérise la loi de probabilité de T .

- La fonction de survie \bar{F} est définie par $\bar{F}(t) = 1 - F(t) = P(T > t)$. C'est la probabilité que le composant fonctionne encore à la date t . \bar{F} est une fonction décroissante, égale à 1 en zero et 0 en l'infini.
- La densité f de la loi de T , lorsqu'elle existe, est caractérisée par

$$F(t) = \int_0^t f(s) ds.$$

Si f est continue, F est alors dérivable de dérivée f . La densité permet de calculer la probabilité que la variable T se réalise dans une partie quelconque⁽²⁾ A de \mathbf{R}_+ :

$$P(T \in A) = \int_A f(s) ds.$$

(1) appelé tribu en théorie de la mesure, contenant l'ensemble vide et Ω , stable par intersection et réunion dénombrables.

(2) d'une certaine tribu, cf. note précédente.

– L'espérance mathématique de la variable aléatoire T est

$$E[T] = \int_0^{\infty} t f(t) dt = \int_0^{\infty} \bar{F}(t) dt.$$

L'égalité des deux derniers termes s'obtient par intégration par parties. Cette quantité est la moyenne théorique de la variable T . Elle peut être finie ou infinie. Pour modéliser des durées de vie, on ne considère généralement que des variables aléatoires d'espérance finie. Si l'on dispose d'observations indépendantes T_1, \dots, T_n , de même loi que T , la loi des grands nombres exprime que la moyenne empirique $\bar{T}_n = (T_1 + \dots + T_n)/n$ converge vers la moyenne théorique $E[T]$ (qu'elle soit finie ou non, dans le cas de variables positives).

– Le taux de panne

$$r(t) = \frac{f(t)}{\bar{F}(t)}.$$

r représente la probabilité « infinitésimale » de panne à l'instant t , sachant que le composant a survécu jusqu'à la date t . Il est facile de voir que si f est continue,

$$r(t) = \lim_{s \rightarrow 0} \frac{P(t < T \leq t + s)}{sP(T > t)} = \lim_{s \rightarrow 0} \frac{F(t + s) - F(t)}{s\bar{F}(t)}.$$

La fonction de survie \bar{F} peut s'exprimer en fonction du taux de panne

$$\bar{F}(t) = e^{-\int_0^t r(s) ds} \quad (1)$$

3. Distributions de durée de vie

Nous allons maintenant étudier quelques exemples de lois de durée de vie. C'est le taux de panne qui est le critère le plus important pour déterminer quelles lois mathématiques peuvent être utilisées pour modéliser pertinemment une durée de vie, puisqu'il mesure l'effet du vieillissement sur un composant individuel. Il déterminera aussi l'intérêt éventuel du brûlage. Si le taux de panne de la distribution F est constant ou croissant, alors le brûlage est inutile ou même néfaste puisqu'il va réduire la durée de vie du composant sans améliorer ses performances. Si au contraire le taux de panne est décroissant sur un intervalle $[0, t_1]$, alors le brûlage devient pertinent. Un taux de panne ayant une réalité physique est nécessairement croissant au delà d'une certaine durée. La relation (1) est donc très utile puisqu'elle permet de caractériser la distribution par son taux de panne. La modélisation sera donc directement basée sur le taux de panne.

On retrouve tout d'abord un premier résultat classique : les seules lois de probabilité à taux de panne constant sont les lois exponentielles, de fonction de survie $\bar{F}(t) = e^{-\lambda t}$, et de densité $f(t) = \lambda e^{-\lambda t}$. Ces lois sont très importantes en modélisation aléatoire. Elles caractérisent les phénomènes dits « sans mémoire ». Dans le contexte présent, elles peuvent être utilisées pour modéliser la durée de vie de composants ne s'usant pas. Dans le folklore de la fiabilité, on l'utilise souvent

pour modéliser la durée de vie d'ampoules. Les autres lois fréquemment utilisées en fiabilité sont les lois de Pareto et les lois de Weibull.

– La loi de Pareto $\mathcal{P}(\alpha)$:

$$\bar{F}(t) = (1+t)^{-\alpha}, \quad r(t) = \alpha(1+t)^{-1}.$$

– La loi de Weibull $\mathcal{W}(c, \theta)$:

$$\bar{F}(t) = e^{-ct^{-\theta}}, \quad r(t) = c\theta t^{\theta-1}.$$

Le paramètre c est un paramètre d'échelle qui n'apporte pas d'information importante sur la loi. Le paramètre θ est appelé paramètre de forme et caractérise la nature de la loi. Pour $\theta = 1$, on retrouve la famille des lois exponentielles.

– La loi de Raleigh est la loi de Weibull de paramètre de forme égal à 2 traduite d'un paramètre θ :

$$\bar{F}(\theta, t) = e^{-\theta t - t^2}, \quad r(\theta, t) = \theta + 2t.$$

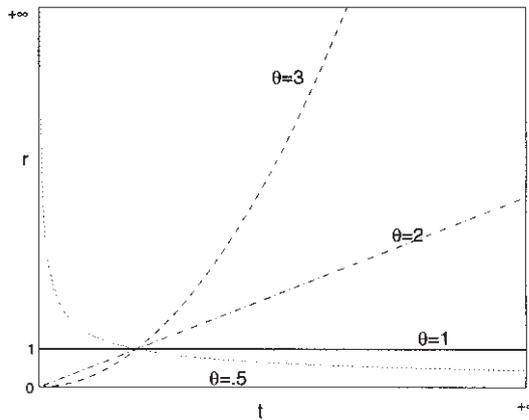


Figure 1 - Taux de panne de la loi de Weibull.

On dira qu'une distribution de probabilité F est à taux de panne « en baignoire » s'il existe des nombres réels positifs $0 \leq t_1 \leq t_2 \leq \infty$, appelés points critiques, tels que

$$r(t) \begin{cases} \text{décroît pour } 0 \leq t \leq t_1, \\ \text{est constant pour } t_1 \leq t \leq t_2, \\ \text{croît pour } t_2 \leq t \leq \infty. \end{cases}$$

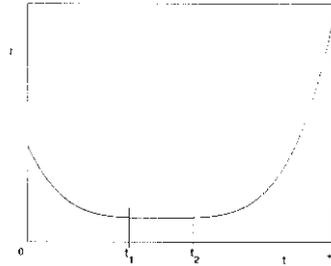


Figure 2 - Taux de panne en baignoire.

3.1 Mélange de populations

Les exemples précédents montrent que les lois de probabilité théoriques simples présentent des taux de panne monotones et ne peuvent pas toutes être utilisées pour justifier la pratique du brûlage. Dans les systèmes industriels, il est naturel de concevoir la durée de vie d'un composant comme un mélange de plusieurs durées de vie, pour rendre compte des variations dans le processus de fabrication.

La notion de mélange se traduit en théorie des probabilités par la caractérisation de la durée de survie d'un composant par un paramètre θ qui est la réalisation d'une variable aléatoire Θ . Pour simplifier la présentation, on se restreindra au cas où Θ prend un nombre fini de valeurs $\theta_1, \dots, \theta_n$ avec

$$P(\Theta = \theta_i) = p_i > 0, \quad \sum_{i=1}^n p_i = 1.$$

Si le composant a la caractéristique θ_i , la loi de sa durée de survie admet la densité $f(\theta_i, t)$. Par analogie avec la formule des probabilités totales, on voit que la densité de la loi de la durée de vie du mélange et sa fonction de survie sont définies par

$$f(t) = \sum_{i=1}^n p_i f(\theta_i, t), \quad \bar{F}(t) = \sum_{i=1}^n p_i \bar{F}(\theta_i, t).$$

Le taux de panne du mélange s'écrit alors :

$$r(t) = \frac{f(t)}{\bar{F}(t)} = \frac{\sum_{i=1}^n p_i f(\theta_i, t)}{\sum_{i=1}^n p_i \bar{F}(\theta_i, t)}. \quad (2)$$

On remarquera que le taux de panne du mélange ne s'exprime pas en fonction des taux de panne individuels de la même façon que la densité ou la fonction de survie. Ceci explique que le taux de panne du mélange peut avoir des propriétés très différentes des taux de pannes des composants individuels, comme le montrent les exemples suivants.

Mélange de deux lois exponentielles. On tire au hasard des composants dont la durée de vie est exponentielle de paramètre 1 ou 2. La probabilité que le composant ait une durée de vie exponentielle de paramètre 1 est p . Le taux de panne du mélange est alors

$$r(t) = \frac{p e^{-t} + 2(1-p)e^{-2t}}{p e^{-t} + (1-p)e^{-2t}}.$$

On voit que dans les deux cas ($p = 1$ et $p = 2$), le taux de panne du mélange est décroissant. Le brûlage est inutile pour une loi exponentielle mais devient pertinent pour un mélange.

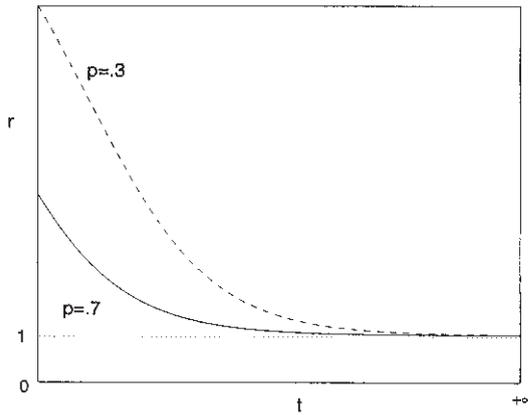


Figure 3 - Taux de panne du mélange de deux lois exponentielles.

Mélange de lois de Raleigh. On tire au hasard parmi des lois ayant des taux de pannes $\theta_i + 2t$, $i = 1, \dots, k$, avec probabilités respectives p_i . Le taux de panne du mélange est donc :

$$r(t) = \frac{\sum_{i=1}^k p_i (\theta_i + 2t) e^{-\theta_i t - 2t^2}}{\sum_{i=1}^k p_i e^{-\theta_i t - 2t^2}} = \frac{\sum_{i=1}^k p_i \theta_i e^{-\theta_i t}}{\sum_{i=1}^k p_i e^{-\theta_i t}} + 2t.$$

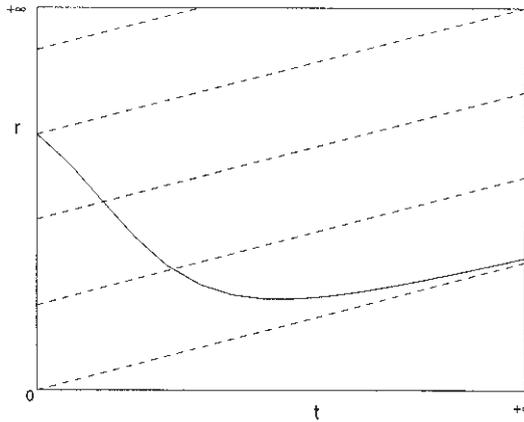


Figure 4 - Taux de panne du mélange de lois de Rayleigh.

Taux de panne limite pour un mélange. On remarque sur les graphes précédents un phénomène conforme à l'intuition : le taux de panne du mélange semble converger vers le taux de panne de l'élément de meilleure qualité, c'est-à-dire ayant le taux de panne le plus faible. Dans les exemples traités précédemment, on peut démontrer cette propriété directement, et elle est vraie dans un contexte plus général. Dans le cas du mélange de lois de Rayleigh, on voit que le taux de panne du mélange a pour asymptote la droite $\theta^* + 2t$ avec $\theta^* = \inf_i \theta_i$.

4. Critères

On peut définir deux types de critères permettant de choisir une durée optimale de brûlage. La durée de brûlage peut être choisie de façon à optimiser la probabilité pour chaque composant d'assurer un temps de service imparti τ , ou la durée de vie moyenne de chaque composant. Ce type de critère est appelé critère de performance puisqu'il s'agit d'optimiser la performance individuelle de chaque composant ou système. Le deuxième type de critère est lié à une fonction de coût qui compense le coût du brûlage par un gain lié à la réalisation d'une mission ou à la durée de vie moyenne des composants ayant survécu au brûlage. Un exemple de chaque type est décrit ci-dessous.

4.1. Critère de performance

Pour un temps de service τ fixé à l'avance, on cherche la durée de brûlage qui maximise la probabilité d'assurer ce temps de service r .

$$b^* = \inf \arg \max_{b \geq 0} \frac{\bar{F}(b + \tau)}{\bar{F}(b)}.$$

Soit F une distribution de durée de vie admettant un taux de panne en baignoire avec points critiques t_1 et t_2 et soit τ une durée de service. On peut alors déterminer la durée optimale théorique du brûlage b^* pour le critère de performance.

Si $\tau \leq t_2 - t_1$, $b^* = t_1$.

Si $\tau > t_2 - t_1$, $b^* < t_1$.

Si $\tau_1 < \tau_2$, alors $b_1^* > b_2^*$, où b_i^* est la durée optimale de brûlage relative à la durée de service τ_i ($i = 1, 2$).

4.2. Critère de coût

On détermine ici la durée optimale de brûlage par un critère de coût. Pour simplifier, on considère que le brûlage a un coût proportionnel à sa durée, avec un facteur de proportionnalité c_0 . Chaque composant n'ayant pas survécu au brûlage coûte c_s . Notons $h(b)$ le coût (aléatoire) du brûlage, c'est-à-dire ce que coûte le brûlage d'un certain nombre (aléatoire) de composants jusqu'à obtention d'un composant ayant survécu. Si l'on représente les composants testés successivement par des variables aléatoires T_1, T_2, \dots et si b est la durée de brûlage, le nombre de composants testés est $\eta = \inf \{i \geq 1, T_i \geq b\}$. On a alors :

$$h(b) = c_0 \sum_{i=1}^{\eta-1} T_i + c_0 b + c_s (\eta - 1) \quad (3)$$

La mise en service d'un composant (ayant survécu au brûlage) assure un gain supposé proportionnel à sa durée de fonctionnement : $G(b) = K(T - b)$ si $T > b$ et $G(b) = 0$ sinon.

Rappelons que l'on suppose que l'espérance, notée μ , de T est finie, c'est-à-dire $\mu = \int_0^{\infty} \bar{F}(s) ds < \infty$. Pour que le brûlage ait un sens, il est nécessaire que $c_s < K\mu$, puisque $K\mu$ représente le gain espéré d'un composant mis en service neuf.

Le coût total (aléatoire) du brûlage est donc $h(b) - G(b)$ (un coût négatif étant un gain). La fonction de coût est l'espérance de ce coût total aléatoire : $c(b) = E[h(b) - G(b)]$. On peut démontrer, en utilisant la théorie des martingales, que c s'exprime de la façon suivante.

$$c(b) = \frac{c_0 \int_0^b \bar{F}(t) dt}{\bar{F}(b)} + \frac{c_s F(b)}{\bar{F}(b)} - K \frac{\int_b^{\infty} \bar{F}(s) ds}{\bar{F}(b)}.$$

4.2.1. Estimation de la durée optimale de brûlage

La durée optimale du brûlage est donc le nombre b^* qui minimise la fonction de coût c . On peut estimer b^* par une méthode graphique très simple. Commençons par donner une autre expression à la fonction de coût. Puisque l'on a supposé

$$\mu = \int_0^{\infty} \bar{F}(s) ds < \infty, \text{ on peut écrire}$$

$$c(b) = \frac{(c_0 + K) \int_0^b \bar{F}(s) ds + c_s - K\mu}{\bar{F}(b)} - c_s.$$

Supposons pour simplifier que f soit strictement positive. F est alors strictement croissante et bijective de $[0, \infty[$ sur $[0, 1[$. On peut donc définir la fonction ϕ de $[0, 1]$ dans \mathbf{R} par :

$$\phi(u) = \mu^{-1} \int_0^{F^{-1}(u)} \bar{F}(s) ds.$$

En posant de plus $\alpha = (K - c_s/M)/(c_0 + K)$, la fonction de coût s'écrit alors

$$c(b) = (c_0 + K)\mu \frac{\phi(F(b)) - \alpha}{1 - F(b)} - c_s.$$

Minimiser $c(b)$ est donc équivalent à maximiser

$$M(u) = \frac{\alpha - \phi(u)}{1 - u},$$

c'est-à-dire que si u^* maximise M , alors $b^* = F^{-1}(u^*)$ minimise c . Ce résultat a une interprétation graphique très simple : la durée optimale de brûlage est $b^* = F(u^*)$ où u^* maximise la pente reliant les points $(u, \phi(u))$ et $(1, \alpha)$.

La fonction ϕ et le paramètre α étant inconnus, il est nécessaire de les estimer à partir d'observations. Supposons que l'on dispose de n observations T_1, \dots, T_n correspondant à un test effectué jusqu'à usure complète de n composants.

Pour estimer la fonction de répartition inconnue F , on utilise la fonction de répartition empirique de l'échantillon T_1, \dots, T_n , définie par :

$$F_n(t) = n^{-1} \text{card}\{i : T_i \leq t\}.$$

Un résultat connu sous le nom de Théorème de Glivenko-Cantelli assure que F_n converge uniformément vers F .

En posant $\bar{F}_n(t) = 1 - F_n(t)$ et $\mu_n = n^{-1} \sum_{i=1}^n T_i$ (la moyenne empirique) on obtient un estimateur empirique de ϕ :

$$\phi_n(u) = \mu_n^{-1} \int_0^{F_n^{-1}(u)} \bar{F}_n(s) ds.$$

ϕ_n converge uniformément vers ϕ sur $[0, 1]$.

On peut calculer aisément les valeurs de la fonction ϕ_n aux points i/n . En supposant l'échantillon ordonné, c'est-à-dire $T_1 \leq T_2 \leq \dots \leq T_n$, et en posant $T_0 = 0$, on voit que $\phi_n(i/n) = U_i/U_n$, avec

$$U_i = \int_0^{F_n^{-1}(i/n)} \bar{F}_n(s) ds = n^{-1} \sum_{j=1}^i (n - j + 1)(T_j - T_{j-1}), \quad i = 1, \dots, n.$$

Finalement, on définit $\alpha_n = (K - c_s/M_n)/(K + c_0)$, qui converge vers α . Il est donc naturel d'estimer u^* par i^*/n où i^* maximise la pente entre les points $(i/n, U_i/U_n)$ et $(1, \alpha)$. La durée optimale de brûlage est donc estimée par $b_n^* = F_n^{-1}(i^*/n)$ qui converge vers b^* lorsque la taille de l'échantillon tend vers l'infini.

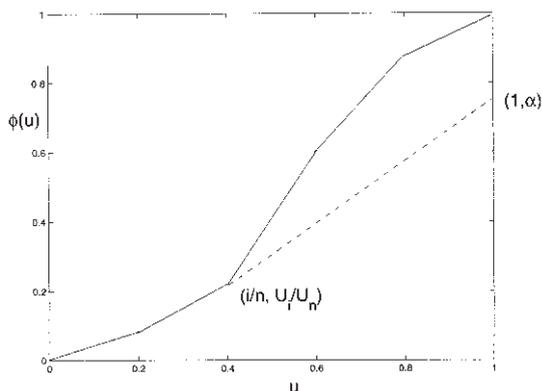


Figure 5 - Méthode graphique d'estimation de la durée optimale du brûlage.

5. Conclusion

Ce texte a tenté de montrer comment formaliser mathématiquement une technique industrielle empirique, physiquement peu fondée. La théorie des probabilités donne un cadre adéquat à ce problème en le ramenant à un problème d'échantillonnage au sein de populations de caractéristiques différentes. Dans la section 3, on a montré que les propriétés théoriques de la modélisation correspondent bien à l'intuition. Ceci justifie l'application de résultats théoriques obtenus à partir de la modélisation, par exemple la méthode d'estimation de la durée optimale proposée dans la section 4.2.1.

Références bibliographiques

Ce texte est une version simplifiée du texte « Le brûlage », sujet de l'épreuve orale de modélisation de l'agrégation externe de mathématiques en 1999 et 2000, publié dans « Mathématiques en situation », C. Ruget (Éd.), collection Scopos, volume 11, Springer (2000). Il est adapté de l'article « Burn-in » de H.W. Block et T.H. Savits, paru en 1997 dans *Statistical Science*⁽³⁾, volume 12, numéro 1, pages 1-19, et du commentaire critique qui le suit.

(3) Cette revue éditée par l'Institute of Mathematical Statistics présente des articles historiques, méthodologiques, de synthèse sur des thèmes nouveaux, polémiques sur l'usage des statistiques en sciences sociales ou en médecine et des entretiens avec quelques-uns des plus grands probabilistes ou statisticiens. Elle est très abordable et intéressante pour toute personne ayant une formation de base en probabilités et statistiques.